

Coursera - IBM Data Science Professional Certification Program

Applied Data Science Capstone Course - Week 1

Project: Battle of Neighborhoods – New Restaurant in Central Ohio, US

By Venkatachalapathy Punugu

1. Introduction

A. Business Problem

Ohio is a state in the East North Central Region of the Midwestern United States – its 17th and incorporated into union in 1803. Ohio has given **8 Presidents (Presidents - WH Harrison, Grant, Hayes, Garfield, B Harrison, McKinley, Taft, Harding)**, Notable Astronauts like **John Glenn** (First American in Space), **Neil Armstrong** (First Man to walk on Moon – Apollo 11), **John Lovell** (Famous for bringing back live – Apollo 13), **Judith Resnick** (Part of Space Shuttle Challenger Crew exploded shortly after takeoff). Franklin County (Names after **Benjamin Franklin**, Statesman, Scientist, and Inventor) is in central Ohio and is the jurisdiction which houses Capitol City of Columbus which is the seat of power for the State of Ohio. Columbus is also the largest City in Ohio.

With its business-friendly atmosphere, large vibrant population, growing diverse demography, Columbus and its immediate neighborhood cities in Franklin County offer many business opportunities for new businesses. Franklin County also houses many of the Top 20 cities in terms of Median Family Income, one of the important indicators of disposable income which will influence the spend on the consumption especially on food and restaurants. It is also observed that Asian descent and especially Indian descent population is one of the top earning demography and therefore an opportunity for the success of the business with a focus on their culture and tastes.

The present exercise is to study Columbus and its neighborhood of Franklin County with a view to recommend three cities for establishing an Indian Restaurant in one of them.

B. Target Audience:

- a. Entrepreneurs targeting to open Indian Restaurant in the state.
- b. Any other interested party or established entities in restaurant business in opening a restaurant in Central Ohio.
- c. Data Science Students to carry forward the analysis and bring out new insights with the help of additional data/analysis/research.

2. Data

- a. **City Data:** City of Columbus lies within the Franklin County and is surrounded by several cities in suburbs. These cities also fall within the jurisdiction of Franklin County. Some of these cities are also target for the location of the restaurant. Therefore, the list of cities in the Franklin County is used for the analysis. [Error! Reference source not found.]
- b. **Location: ARCGIS** is used to retrieve the latitude and longitude data and we will be using the geocode function for the purpose.
- c. **Venues in the cities: FOURSQUARE API** is used to retrieve all the venues within the cities in the analysis
- d. A radius of 500 meters were used to retrieve the venues in each of the cities – these however were giving a very short list and at the same time, based on the latitude and longitude of the center of the target city, resulted in not picking up the venues and city. Some important cities like Dublin were totally missing. Therefore, the radius was increased to **2000 meters**. This provided a good selection of the venues across all the cities in the county.
- e. **Income Data:** We will be using a commercial data that is derived from the US Census to retrieve the median income of the cities. [Error! Reference source not found.]
- f. **SKLEARN** – Scikit Learn package is used for arriving at the KMean for the venues.
- g. **Visualization:**
 - a. **Folium** package is used for the geographical mapping and cluster visualization.
 - b. **Yellowbrick** is used for the visualization using the KElbowVisualizer for displaying the optimum number of K-Mean clusters to be used for grouping the venues.
 - c. **Matplotlib.pyplot** is used for the visualization of the venues across the cities
- h. **Data Cleanup:** As needed based on the availability of the above data.
- i. **Data Usage:** The above data will be used to arrive at top 3 cities for the recommendation.